Math. Model. Nat. Phenom. 21 (2026) 19
https://doi.org/10.1051/mmnp/2024014

Mathematical Modelling of Natural Phenomena
www.mmnp-journal.org

DATA-DRIVEN NEURONAL DYNAMICS: MODEL IDENTIFICATION AND SPIKING ANALYSIS

Yicheng Mao[1] and Xianbin Liu[1,*]

Abstract. Neurons serve as the fundamental unit in the structure and functionality of the nervous system, and the large cluster network formed by their interconnection determines the perception, learning, emotion, memory and other behaviors of the organism. While prior studies of neurons primarily focused on costly experiments and feature-based modeling, recent advancements in computing power and algorithms have made it possible to identify and analyze system with data-driven approaches. This study introduces a data-driven method for the identification of neuronal systems, and the result shows that the method is of great precision for further study. In addition, in order to investigate the mechanisms underlying the transition between resting and active states, a finite difference scheme is developed to calculate the mean first exit time and spiking frequency of neuronal systems. Notably, we propose the concept of noise coupling effect in our study, which elaborates that the spiking of a neuronal dynamical system can occur more rapidly under less stringent conditions when driven by the potential fluctuation of both Gaussian white noise and Lévy noise.

Received April 23, 2024. Accepted June 24, 2024.

1. Introduction

The study of neuronal dynamics has emerged as a critical field within brain science, garnering increasing importance and witnessing significant advancements over the years. Neuronal dynamics refers to the patterns and mechanisms of activity that occur within and between neurons, shaping the complex functioning of the brain. Understanding neuronal dynamics is crucial for unraveling the mysteries of the brain and comprehending various cognitive processes, such as perception, learning, and memory.

The intricate network of neurons in the brain communicates through electrical and chemical signals, forming the basis for neural information processing. The investigation of neuronal dynamics aims to decipher the principles underlying these neural interactions and shed light on how they give rise to cognitive function. By studying the dynamics of individual neurons and analyzing their collective behaviors, we seek to gain insight into the fundamental principles of brain function.

Advancements in experimental techniques have played a pivotal role in driving the development of research on neuronal dynamics. Over the past decades, researchers have developed sophisticated mathematical models that simulate the behavior of neuronal networks, allowing for the investigation of emergent behaviors and

Keywords and phrases: Nonlinear stochastic dynamics, neuronal dynamic system, data-driven model identification, the exit problem in stochastic dynamics.

[1] Nanjing University of Aeronautics and Astronautics, Nanjing, China.

[*] Corresponding author: xbliu@nuaa.edu.cn

collective phenomena. A concise yet robust classification of excitable neurons was put forward by Hodgkin [1], where the neuronal excitability is classified into two types: Class 1 excitability implies that action potentials of any frequency can be generated under electric stimulation, while Class 2 excitability means that action potentials can only be generated in a certain frequency band. Scholars have proposed the bifurcation theory of neuronal dynamics to provide further insights into the firing types of neurons. Rinzel & Ermentrout [2] employed the current intensity as a bifurcation parameter to establish a well-known paradigm for neuron models. Class 1 excitability, arising from a saddle-node bifurcation [3–8], and Class 2 excitability, observed when a rest potential loses stability *via* an Andronov–Hopf bifurcation [6, 8–13], are two examples of bifurcation whose codimension is equal to 1. Moreover, scholars have also extensively classified bifurcations for complex cases with higher codimension [6, 8, 14–18].

With the rapid development of modern computing technology and scientific tools, various fields have started to embrace the new paradigm of data-driven development, which includes medical imaging diagnosis, autonomous driving technology, financial risk control strategies, and even open AI assistants [19–21]. Thanks to the modern deluge of data, data-driven system identification has developed rapidly in recent years. For deterministic systems, Brunton et al. [22, 23] have developed tools of sparse identification to identify ordinary differential equations. Furthermore, in the context of stochastic systems, data-driven identification methods can be divided into three main categories. The first way is to directly identify the drift and diffusion coefficients, which includes the methods of neural differential equations [24, 25], Bayesian inference [26, 27], and non-local Kramers–Moyal formulae [28–30]. The second approach focuses on learning the generator of stochastic differential equations by estimating the stochastic Koopman operator [31–33]. Lastly, In the third way, the coefficients are obtained by estimating the Fokker–Planck equations corresponding to stochastic differential equations [34, 35].

The thesis of this paper is twofold. Firstly, we propose a data-driven method for identifying neuron models, particularly those driven by non-Gaussian Lévy potential fluctuations. Secondly, we examine the impact of various potential stochastic fluctuations on the spike period. This paper is organized as follows. In Section 2, we introduce the paradigm of Ermentrout–Kopell neuron models [3, 36], and then we utilize non-local Kramers–Moyal formulae [28–30] to realize the model identification. In Section 3, by numerically solving the mean first exit time, we estimate the variation of the spike period under different non-Gaussian Lévy potential fluctuations, and we propose the noise coupling effect in our analysis, which enables us to study the nature of spontaneous periodic firing in neurons. Conclusions are presented in Section 4. Additionally, for the sake of clarity and coherence, specific details are arranged in the appendix for reference.

2. Data-driven neuronal model identification

The neuronal excitability is marked by significant fluctuations in action potentials. These fluctuations are produced and maintained by the movement of ions through the cell membrane. In the simplest scenario, an elevated membrane potential stimulates the opening of Na+ and Ca++ channels, causing a swift influx of these ions and consequently raising the membrane potential even further. Subsequently, as a positive feedback mechanism regulates this sudden increase, the channels become inactivated, and K+ channels are activated. As a result, the K+ current increases, leading to a reduction in the membrane potential. This process is commonly known as a spike in neuronal dynamics, and Figure 1 depicts the spikes triggered by the initial potential fluctuation.

Experimental observations have consistently shown that the spontaneous firing of neurons is often attributed to random fluctuations in their initial potential. Under the assumption that the clustering behavior of neuronal Poisson spikes can be approximated using the central limit theorem, previous studies have primarily focused on investigating this pure-diffusion phenomenon under the perturbation of additive Gaussian white noise [37–40]. Nevertheless, it has been observed that systems comprising numerous nonlinear subneurons often display non-Gaussian discontinuous spikes. Therefore, exploring neuronal dynamics under non-Gaussian Lévy noise provides a more accurate representation of these abrupt potential jumps [41–45]. In this paper, we further examine one

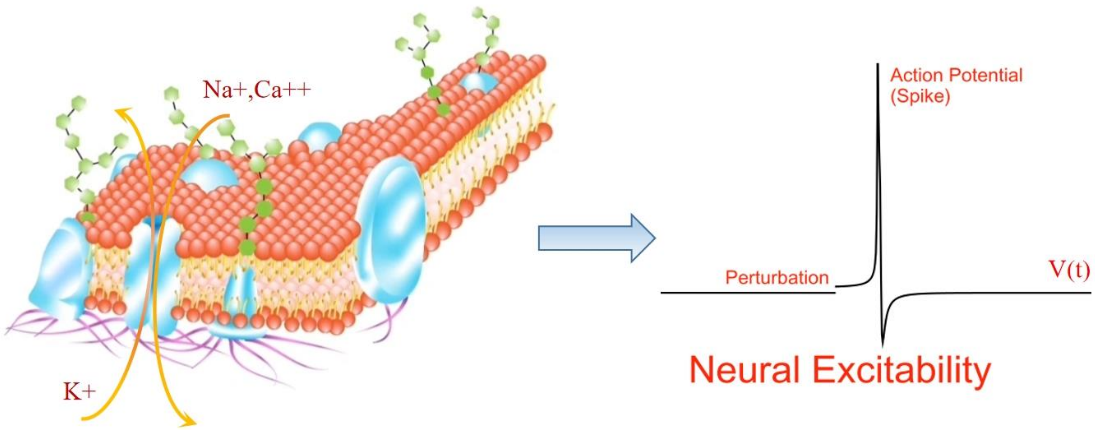


FIGURE 1. Diagram of neural spikes caused by initial potential fluctuations

of the most general scenarios, where multiplicative white Gaussian noise and additive Lévy noise act together as random perturbations on the self-excited discharge of neurons.

2.1. Non-local Kramers–Moyal formula for system identification

In order to identify the stochastic dynamical system driven by Lévy noise and multiplicative Gaussian white noise, based on the FPK equation of the solution process, Li and Duan [28–30] derived the following non-local Kramers–Moyal formula to extract the drift terms, diffusion terms and noise parameters from the sample path.

Consider an n-dimensional stochastic dynamical in the following form

$$\mathrm{d}\boldsymbol{X}(t)=\boldsymbol{b}(\boldsymbol{X}(t))dt+\boldsymbol{\Lambda}(\boldsymbol{X}(t))\mathrm{d}\boldsymbol{B_t}+\sigma\mathrm{d}\boldsymbol{L_t} \tag{2.1}$$

where $\boldsymbol{B}_t=[B_{1,t},B_{2,t},...,B_{n,t}]$ is an n-dimensional Brownian motion. $\boldsymbol{L}_t=[L_{1,t},L_{2,t},...,L_{n,t}]$ denotes an n-dimensional non-Gaussian Lévy motion with triplet $(0,0,\nu_\alpha)$, where the jump measure of L_t satisfies $\nu(\mathrm{d}\boldsymbol{y})=W(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$ for $\boldsymbol{y}\in R^n/\{0\}$, specific details can be found in Definition A.3 of Appendix A. The vector $\boldsymbol{b}(\boldsymbol{x})=[b_1(\boldsymbol{x}),b_2(\boldsymbol{x}),...,b_n(\boldsymbol{x})]^{\mathrm{T}}$ represents the drift coefficient, and the diffusion coefficient $\boldsymbol{\Lambda}(\boldsymbol{x})$ is an $n\times n$ matrix. The covariance matrix is defined by $\boldsymbol{a}(\boldsymbol{x})=\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\mathrm{T}}$.

Theorem 2.1 (non-local Kramers–Moyal formula [28–30])**.** *For every $\varepsilon>0$, the probability density $p(\boldsymbol{x},t|\boldsymbol{z},0)$, where $\boldsymbol{z}=\boldsymbol{x}(0)$, has the following relations with the jump measure, drift and diffusion term.*

(i) For every $\boldsymbol{x}$ and $\boldsymbol{z}$ satisfying $|\boldsymbol{x}-\boldsymbol{z}|>\varepsilon$, $\lim\limits_{t\to 0}p(\boldsymbol{x},t|\boldsymbol{z},0)/t=\sigma^{-n}W(\sigma^{-1}(\boldsymbol{x}-\boldsymbol{z}))$ uniformly in $\boldsymbol{x}$ and $\boldsymbol{z}$;
(ii) For $i=1,2,...,n$, $\lim\limits_{t\to 0}t^{-1}\int_{|\boldsymbol{x}-\boldsymbol{z}|<\varepsilon}(x_i-z_i)p(\boldsymbol{x},t|\boldsymbol{z},0)\mathrm{d}\boldsymbol{x}=b_i(\boldsymbol{z})$;
(iii) For $i,j=1,2,...,n$, $\lim\limits_{t\to 0}t^{-1}\int_{|\boldsymbol{x}-\boldsymbol{z}|<\varepsilon}(x_i-z_i)(x_j-z_j)p(\boldsymbol{x},t|\boldsymbol{z},0)\mathrm{d}\boldsymbol{x}=a_{ij}(\boldsymbol{z})+\sigma^{-n}\int_{|\boldsymbol{y}|<\varepsilon}y_iy_jW(\sigma^{-1}\boldsymbol{y})\mathrm{d}\boldsymbol{y}$.

Corollary 2.2. *For every $\varepsilon>0$, the solution process $\boldsymbol{X}(t)$ of (1) has the following relations with the jump measure, drift and diffusion term.*

(i) For every $m>1, m\in Z^+$, $\lim\limits_{t\to 0}t^{-1}P\{|x(t)-z|\in[\varepsilon,m\varepsilon)|x(0)=z\}=\sigma^{-n}\int_{|y|\in[\varepsilon,m\varepsilon)}W(\sigma^{-1}y)\mathrm{d}y$;
(ii) For $i=1,2,...,n$, $\lim\limits_{t\to 0}t^{-1}P\{|\boldsymbol{x}(t)-\boldsymbol{z}|<\varepsilon|\boldsymbol{x}(0)=\boldsymbol{z}\}\cdot E[(x_i(t)-z_i)|\boldsymbol{x}(0)=\boldsymbol{z};|\boldsymbol{x}(t)-\boldsymbol{z}|<\varepsilon]=b_i(\boldsymbol{z})$;
(iii) For $i,j=1,2,...,n$, $\lim\limits_{t\to 0}t^{-1}P\{|\boldsymbol{x}(t)-\boldsymbol{z}|<\varepsilon|\boldsymbol{x}(0)=\boldsymbol{z}\}\cdot E[(x_i(t)-z_i)(x_j(t)-z_j)|\boldsymbol{x}(0)=\boldsymbol{z};|\boldsymbol{x}(t)-\boldsymbol{z}|<\varepsilon]=a_{ij}(\boldsymbol{z})+\sigma^{-n}\int_{|\boldsymbol{y}|<\varepsilon}y_iy_jW(\sigma^{-1}\boldsymbol{y})\mathrm{d}\boldsymbol{y}$.

2.2. Identification for the Lévy motion

Utilizing Corollary A.5, we can identify the Lévy motion, the drift term and covariance matrix from sample paths.

Assume that there exists a pair of datasets observed from equation (2.1)

$$\begin{aligned} \boldsymbol{Z} &= [\boldsymbol{z}_1, \boldsymbol{z}_2, ..., \boldsymbol{z}_M], \\ \boldsymbol{X} &= [\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_M]. \end{aligned} \tag{2.2}$$

where $\boldsymbol{x}_i$ is the image of $\boldsymbol{z}_i$ after a small time interval h.

In order to measure the displacement of each point after an evolution time h, we construct a new data set

$$\boldsymbol{R} = [|\boldsymbol{y}_1|, |\boldsymbol{y}_2|, ..., |\boldsymbol{y}_M|], \boldsymbol{y}_i = \boldsymbol{x}_i - \boldsymbol{z}_i. \tag{2.3}$$

Furthermore, a dictionary of basis functions, $\boldsymbol{\varphi}(\boldsymbol{x}) = [\varphi_1(x), \varphi_2(x), ..., \varphi_k(x)]$, is also needed for the identification of the drift term and the covariance matrix.

For the numerical identification of the Lévy noise intensity σ and the stability parameter α, we consider $N+1$ intervals $[\varepsilon, m\varepsilon), [m\varepsilon, m^2\varepsilon), ..., [m^N\varepsilon, m^{N+1}\varepsilon)$for $\varepsilon > 0$ and $m > 1$. Then we counted the number of elements in $\boldsymbol{R}$ that fall within the above intervals, which is denoted as $n_1, n_2, ..., n_{N+1}$. Therefore, based on the first assertion of Corollary A.5, the conditional probability on the left side of the first assertion can be approximated by the ratio of the number of the points falling into the interval $[\varepsilon, m\varepsilon)$ to the total number M, *i.e.*

$$\lim_{h\to 0} P\{|\boldsymbol{x}(h) - \boldsymbol{z}| \in [m^k\varepsilon, m^{k+1}\varepsilon)|\boldsymbol{x}(0) = \boldsymbol{z}\} \approx h^{-1}M^{-1}n_k \tag{2.4}$$

In the system identification of equation (2.1), the integration in the right hand of assertion (i) yields

$$\sigma^{-1}\int_{|y|\in[m^k\varepsilon, m^{k+1}\varepsilon)} W(\sigma^{-1}y)\mathrm{d}y = 2\sigma^{\alpha}c(1,\alpha)\alpha^{-1}\varepsilon^{-\alpha}m^{-k\alpha}(1-m^{-\alpha}) \tag{2.5}$$

for α-stable Lévy noise, $W(y) = c(n,\alpha)|y|^{-(n+\alpha)}$, $c(n,\alpha) = \frac{\alpha\Gamma((n+\alpha)/2)}{2^{1-\alpha}\pi^{n/2}\Gamma(1-\alpha/2)}$.

Combine equation (2.4) and equation (2.5), assertion (i) yields $N+1$ equalities

$$2\sigma^{\alpha}c(1,\alpha)\alpha^{-1}\varepsilon^{-\alpha}m^{-k\alpha} = h^{-1}M^{-1}n_k, k = 0, 1, ..., N. \tag{2.6}$$

The ratios of the first equation ($k=0$) to the other N equations lead to the solutions of α

$$\alpha = (k\ln m)^{-1}\ln\frac{n_0}{n_k}, k = 1, 2, ..., N. \tag{2.7}$$

Therefore, we take the arithmetic mean of equation (2.7) as the identification value of α, which is denoted as $\tilde{\alpha}$. Similarly, substituting $\tilde{\alpha}$ into equation (2.6), we get the solutions of σ, and we finally take the arithmetic mean $\tilde{\sigma}$ as the identification value of σ.

$$\sigma = [\frac{\tilde{\alpha}\varepsilon^{\tilde{\alpha}}m^{k\tilde{\alpha}}n_k}{2c(1,\tilde{\alpha})hM(1-m^{-\tilde{\alpha}})}]^{1/\tilde{\alpha}}, k = 0, 1, ..., N. \tag{2.8}$$

2.3. Identification for the drift term

Based on the second assertion of Corollary A.5, the identification of the drift term requires the statistical information in the region $|\boldsymbol{x}(t)-\boldsymbol{z}|<\varepsilon$. Therefore, we reconstructed the dataset equation (2.2) as follows

$$\begin{aligned}\bar{\boldsymbol{Z}} &= [\bar{\boldsymbol{z}}_1, \bar{\boldsymbol{z}}_2, ..., \bar{\boldsymbol{z}}_{\bar{M}}],\\ \bar{\boldsymbol{X}} &= [\bar{\boldsymbol{x}}_1, \bar{\boldsymbol{x}}_2, ..., \bar{\boldsymbol{x}}_{\bar{M}}].\end{aligned} \tag{2.9}$$

Compared with the original dataset equations (2.2), and (2.9) retains only the part of equation (2.2) that satisfies $|\boldsymbol{x}_i-\boldsymbol{z}_i|<\varepsilon$, where $\bar{M}$ indicates the number of data pairs that meet this condition. Therefore, the conditional probability $P\{|\boldsymbol{x}(t)-\boldsymbol{z}|<\varepsilon|\boldsymbol{x}(0)=\boldsymbol{z}\}$ can be estimated as $\bar{M}M^{-1}$.

In addition, based on the dictionary of the basis functions, each component of the drift coefficient can be approximated as $b_i(\boldsymbol{x}) \approx \sum_{k=1}^{K} c_{i,k}\varphi_k(\boldsymbol{x}), i=1,2,...,n$. Above all, we derive the following equation for corollary (ii)

$$\begin{aligned} &A\boldsymbol{c}_i = B_i,\\ &A = \begin{bmatrix} \varphi_1(\bar{\boldsymbol{z}}_1) & \cdots & \varphi_K(\bar{\boldsymbol{z}}_1)\\ \vdots & & \vdots\\ \varphi_1(\bar{\boldsymbol{z}}_{\bar{M}}) & \cdots & \varphi_K(\bar{\boldsymbol{z}}_{\bar{M}})\end{bmatrix},\\ &B_i = \bar{M}M^{-1}h^{-1}[\bar{\boldsymbol{x}}_{i,1}-\bar{\boldsymbol{z}}_{i,1}, ..., \bar{\boldsymbol{x}}_{i,M}-\bar{\boldsymbol{z}}_{i,M}]^{\mathrm{T}}.\end{aligned} \tag{2.10}$$

where $\boldsymbol{c}_i=[c_{i,1},c_{i,2},...,c_{i,K}]$ is the coordinate of the i- th component of the drift term $\boldsymbol{b}(\boldsymbol{x})$ on the function basis. In general, equation (2.10) has no solutions. Here we utilize the sparse identification method to learn the drift term, which minimizes $||A\boldsymbol{c}_i-B_i||_2$ to find the optimal solution in the least square sense. Therefore, the coefficient of the function basis can be solved as follows

$$\tilde{\boldsymbol{c}}_i = (A^{\mathrm{T}}A)^{-1}(A^{\mathrm{T}}B_i). \tag{2.11}$$

2.4. Identification for the diffusion term

Finally, based on the third assertion of Corollary A.5, we can learn the covariance matrix $a(\boldsymbol{x})$ of the diffusion term. Similar to the identification of the drift term, each component of the covariance matrix can be approximated as $a_{ij}(\boldsymbol{x}) \approx \sum_{k=1}^{K} d_{ij,k}\varphi_k(\boldsymbol{x}), i,j=1,2,...,n$. Therefore, we derive the equation of corollary (iii) as follows

$$\begin{gathered} A\boldsymbol{d}_{ij} = B_{ij},\\ A = \begin{bmatrix} \varphi_1(\bar{\boldsymbol{z}}_1) & \cdots & \varphi_K(\bar{\boldsymbol{z}}_1)\\ \vdots & & \vdots\\ \varphi_1(\bar{\boldsymbol{z}}_{\bar{M}}) & \cdots & \varphi_K(\bar{\boldsymbol{z}}_{\bar{M}})\end{bmatrix},\\ B_{ij} = \bar{M}M^{-1}h^{-1}[(x_{i,1}-z_{i,1})(x_{j,1}-z_{j,1}), ..., (x_{i,\bar{M}}-z_{i,\bar{M}})(x_{j,\bar{M}}-z_{j,\bar{M}})]^{\mathrm{T}} - S_{ij}(\varepsilon),\\ S_{ij}(\varepsilon) = \tilde{\sigma}^{-n}\int_{|\mathrm{y}|<\varepsilon} y_i y_j W(\tilde{\sigma}^{-1}\mathrm{y})\mathrm{dy}.\end{gathered} \tag{2.12}$$

where $\boldsymbol{d}_{ij}=[d_{ij,1},d_{ij,2},...,d_{ij,k}]^{\mathrm{T}}$ and $S_{ij}(\varepsilon)$ requires the parameters $\tilde{\alpha}$ and $\tilde{\sigma}$ estimated in Section 2.3. Again from the sparse identification method, we get the least square solution

$$\tilde{\boldsymbol{d}}_{ij} = (A^{\mathrm{T}}A)^{-1}(A^{\mathrm{T}}B_{ij}) \tag{2.13}$$

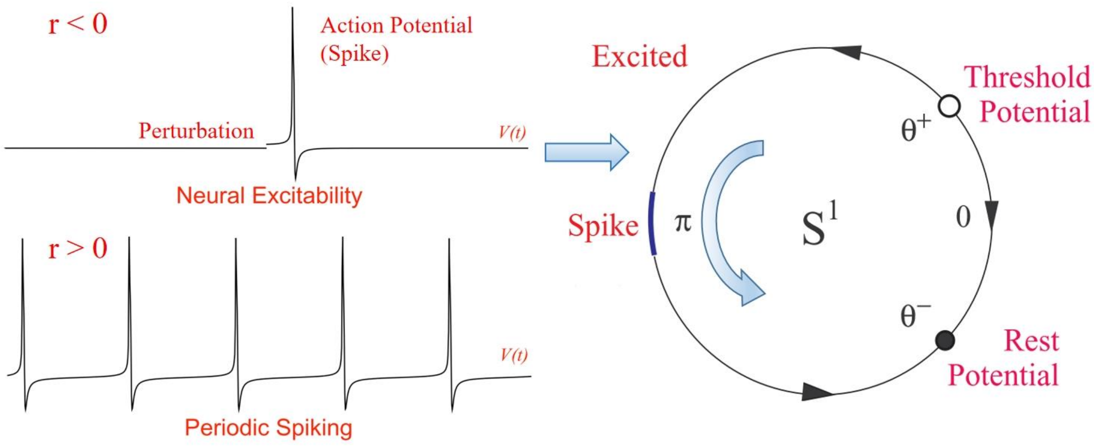


FIGURE 2. Neural excitability ($r < 0$) and Periodic spiking ($r > 0$).

2.5. Excitable neuronal systems driven by Lévy potential excitation

In this section, we mainly examine the Ermentrout–Kopell model [3, 46, 47], which is the normal form of Class 1 excitable neurons:

$$\begin{gathered}\dot{\theta} = (1 - \cos\theta) + (1 + \cos\theta)r, \\ \theta \in S^1, mod\ 2\pi.\end{gathered} \tag{2.14}$$

where $\theta \in [-\pi, \pi)$ represents the membrane potential along the invariant circle, r is the bifurcation parameter. This model reveals the Class 1 excitability through the saddle-node bifurcation on an invariant circle with codimension 1.

If $r<0$, there exists two equilibria for (1), $\theta^{\pm} = \pm\cos^{-1}((1 + r)/(1 - r))$, representing the rest state and the threshold respectively. The system exhibits excitability in the following manner: Small perturbations of θ that do not exceed the threshold θ^+ decay exponentially. Conversely, if θ is perturbed beyond the threshold, it continues to grow and finally generates a spike along the circle. Furthermore, If $r>0$, there are no equilibria for (1), resulting in oscillations of $\theta(t)$ with the frequency of $\omega = 2\sqrt{r}$. Diagrams of each case and neural excitability are shown in Figure 2.

In this section, we validate the data-driven approach by considering a most general case where the neuronal model equation (2.14) is driven by both multiplicative Gaussian white noise and additive pure jump α- stable Lévy noise:

$$\mathrm{d}\theta = (1 - \cos\theta)dt + (1 + \cos\theta)rdt + \sigma_B(1 + \cos\theta)\mathrm{d}B_t + \sigma_L d{L_t}^{\alpha} \tag{2.15}$$

where B_t is Brownian motion with multiplicative diffusion intensity σ_B. ${L_t}^{\alpha}$ is an α- stable Lévy process with triplet $(0, 0, \nu_\alpha)$, and σ_L is the corresponding Lévy noise intensity. Specific details about Lévy process are arranged in Appendix A for reference.

2.6. Extracting SDE from noised data for neuron systems

Utilizing the methods derived in Sections 2.2 to 2.4, we identify the Ermentrout–Kopell neuronal system equation (2.15) in excitable state (r $<$ 0) and periodic spiking state (r $>$ 0) respectively. The results are listed as follows, where the sampling interval $h = 0.01$ and the selected function dictionary is listed as follows.

When identifying the diffusion term, it is not possible for us to directly determine the diffusion coefficient of Gaussian white noise. This limitation arises from the fact that Gaussian white noise is the formal derivative of

TABLE 1. Identified Lévy motion for the excitable state (r < 0 in Fig. 2).

The stability parameter α			Lévy noise intensity σ_L		
True	Learn	Error	True	Learn	Error
1.5	1.4997	2×10^{-4}	2	2.0035	1.75×10^{-3}
1.25	1.2493	5.6×10^{-4}	2	2.0103	5.15×10^{-3}
1	1.0015	1.5×10^{-3}	2	2.0021	1.05×10^{-3}
0.5	0.5018	3.6×10^{-3}	2	1.9912	4.4×10^{-3}

TABLE 2. Identified drift term for the excitable state (r < 0 in Fig. 2).

Basis	$r = -0.1$			$r = -0.2$			$r = -0.3$		
	True	Learn	Error	True	Learn	Error	True	Learn	Error
1	0.9	0.8947	5.9×10^{-3}	0.8	0.7897	1.2×10^{-2}	0.7	0.6975	3.5×10^{-3}
θ	0	0		0	0		0	0	
θ^2	0	0		0	0		0	0	
θ^3	0	0		0	0		0	0	
$\sin\theta$	0	0		0	0		0	0	
$\cos\theta$	–1.1	–1.0911	8.1×10^{-3}	–1.2	–1.1908	7.6×10^{-3}	–1.3	–1.2882	9.1×10^{-3}
$\sin 2\theta$	0	0		0	0		0	0	
$\cos 2\theta$	0	0		0	0		0	0	

TABLE 3. Identified variance of the diffusion term (for the excitable state r < 0 in Fig. 2).

Basis	$\sigma_B = 0.2$			$\sigma_B = 0.5$			$\sigma_B = 1$		
	True	Learn	Error	True	Learn	Error	True	Learn	Error
1	0.04	0.0389	2.7×10^{-2}	0.25	0.2436	2.5×10^{-2}	1	0.9884	1.1×10^{-2}
θ	0	0		0	0		0	0	
θ^2	0	0		0	0		0	0	
θ^3	0	0		0	0		0	0	
θ^4	0	0		0	0		0	0	
$\sin\theta$	0	0		0	0		0	0	
$\cos\theta$	0.08	0.0807	8.7×10^{-3}	0.5	0.4896	2.1×10^{-2}	2	1.9913	4.3×10^{-3}
$\cos^2\theta$	0.04	0.0386	3.5×10^{-2}	0.25	0.2469	1.2×10^{-2}	1	0.9891	1.1×10^{-2}

Brownian motion in the statistical sense, and the increments of Brownian motion follow a normal distribution. Moreover, since the sum of independent normal distributions is also a normal distribution, the same statistical information can correspond to different forms of diffusion coefficients.

Therefore, we calculate the covariance of the diffusion term in Tables 3 and 5, where the basis function dictionary is selected as $\{1, \theta, \theta^2, \theta^3, \theta^4, \sin\theta, \cos\theta, \sin^2\theta, \cos^2\theta\}$. In order to enhance the stability of the identification outcomes, we exclude $\sin^2\theta$ from the set of function bases. The decision was made due to the decomposition of $\sin^2\theta = 1 - \cos^2\theta$, that is, we dismiss the functions that can be seen as linear combinations of other basis functions.

Finally, to more accurately identify the dominant basis function in the function dictionary, we employ the step error verification sparse regression method. The procedure begins by randomly splitting the sample dataset

TABLE 4. Identified drift term for the periodic spiking state (r > 0 in Fig. 2).

Basis	$r = 0.1$			$r = 0.2$			$r = 0.3$		
	True	Learn	Error	True	Learn	Error	True	Learn	Error
1	1.1	1.0864	1.2×10^{-2}	1.2	1.1821	1.5×10^{-2}	1.3	1.2847	1.1×10^{-2}
θ	0	0		0	0		0	0	
θ^2	0	0		0	0		0	0	
θ^3	0	0		0	0		0	0	
$\sin\theta$	0	0		0	0		0	0	
$\cos\theta$	-0.9	-0.9071	7.9×10^{-3}	–0.8	–0.7796	2.5×10^{-2}	–0.7	–0.7097	1.4×10^{-2}
$\sin 2\theta$	0	0		0	0		0	0	
$\cos 2\theta$	0	0		0	0		0	0	

TABLE 5. Identified variance of the diffusion term (for the periodic spiking state r > 0 in Fig. 2).

Basis	$\sigma_B = 0.2$			$\sigma_B = 0.5$			$\sigma_B = 1$		
	True	Learn	Error	True	Learn	Error	True	Learn	Error
1	0.04	0.0379	5.2×10^{-2}	0.25	0. 2444	2.2×10^{-2}	1	0.9771	2.3×10^{-2}
θ	0	0		0	0		0	0	
θ^2	0	0		0	0		0	0	
θ^3	0	0		0	0		0	0	
θ^4	0	0		0	0		0	0	
$\sin\theta$	0	0		0	0		0	0	
$\cos\theta$	0.08	0.0785	1.9×10^{-2}	0.5	0.4869	2.6×10^{-2}	2	1.8979	5.1×10^{-2}
$\cos^2\theta$	0.04	0.0387	3.2×10^{-2}	0.25	0.2471	1.2×10^{-2}	1	0.9863	1.4×10^{-2}

equation (2.9) into two portions: 10% for the test set and 90% for the training set. Subsequently, we conduct sparse regression on the entire function dictionary using the training set, while documenting the mean square error observed in the test set. Next, we eliminate the basis with the smallest absolute value of the regression coefficient, resulting in a new function dictionary. This process is iterated until only one basis remains in the function dictionary. Ultimately, we select the function dictionary corresponding to the minimum mean square error as the optimal sparse regression.

In addition, the above strategy can be applied to system identification for higher-dimensional neuron systems with more complex noise excitation. For instance, for the following FHN neuron system [48–51] driven by Gaussian and Lévy noise:

$$\begin{cases} \mathrm{d}x_1 = (x_1 - {x_1}^3/3 - x_2)dt + \sigma_B(1 + x_2)\mathrm{d}B_{1,t} + \sigma_B \mathrm{d}B_{2,t} + \sigma_L dL^{\alpha}_{1,t} \\ \mathrm{d}x_2 = k(x_1 + m)dt + \sigma_B x_1 \mathrm{d}B_{2,t} + \sigma_L dL^{\alpha}_{2,t} \end{cases} \tag{2.16}$$

where the parameter $k \ll 1$ is the ratio of slow and fast time scales, and m is a control parameter that governs the character of the solutions. The case of $|m| > 1$ was widely used to characterize the phenomenon of nerve pulses [52]. $\boldsymbol{B}_t = [B_{1,t}, B_{2,t}]^{\mathrm{T}}$ is a two-dimensional Brownian motion with independent component, $\boldsymbol{L}^{\alpha}_t = [L^{\alpha}_{1,t}, L^{\alpha}_{2,t}]^{\mathrm{T}}$ is an $\alpha-$stable Lévy process with triplet $(0, 0, \nu_\alpha)$. σ_B and σ_L denote the intensity of Gaussian white noise and Lévy noise.

To maintain the conciseness of the paper, the system identification results for the case $k = 0.2, m = 1.2$ are provided in Appendix C for reference.

3. Neuron spiking analysis

3.1. The exit problem of SDEs driven by Lévy noise

The neuron activity in an excitable state is typically characterized by its average firing frequency. Therefore, in order to reveal the common natural firing characteristics of excitable neurons, we primarily investigate the mean first exit time of the original Class 1 excitable Ermentrout–Kopell neural model, which is driven by both multiplicative Gaussian white noise and additive pure jump α- stable Lévy noise:

$$\begin{aligned}&\mathrm{d}\theta = (1-\cos\theta)dt + (1+\cos\theta)rdt + \sigma_B(1+\cos\theta)\mathrm{d}B_t + \sigma_L dL_t{}^{\alpha},\\ &\theta\in S^1,\ mod 2\pi.\end{aligned} \tag{3.1}$$

where $r = -0.2$ corresponds to the Class 1 excitable state. σ_B and σ_L are the intensity of Gaussian white noise and Lévy noise respectively.

Unlike the case where it is solely stimulated by Gaussian white noise, the solution trajectory of the stochastic dynamical system driven by Lévy noise no longer exhibits random continuity and instead demonstrates jumping characteristics.

For a general form of stochastic dynamical system driven by α- stable Lévy noise:

$$\mathrm{d}\boldsymbol{X_t} = \boldsymbol{b}(\boldsymbol{X_t})dt + \boldsymbol{\sigma_B}(\boldsymbol{X_t})\mathrm{d}\boldsymbol{B_t} + \boldsymbol{\sigma_L}\mathrm{d}\boldsymbol{L_t}^{\boldsymbol{\alpha}} \tag{3.2}$$

where $\boldsymbol{X_t}\in R^n, \boldsymbol{b}(\boldsymbol{X_t})\in R^n, \boldsymbol{B_t}\in R^m, \boldsymbol{\sigma_B}(\boldsymbol{X_t})\in R^{n\times m}, \boldsymbol{\sigma_L} = diag(\sigma_{L,1},\sigma_{L,2},...,\sigma_{L,n})$ is the Lévy noise intensity. $\boldsymbol{L_t}^{\boldsymbol{\alpha}} = ({L_t}^{\alpha_1}, {L_t}^{\alpha_2}, ..., {L_t}^{\alpha_n})^T$, and each component ${L_t}^{\alpha_i}, (i = 1, ..., n)$ is an α-stable Lévy process with triplet $(0, 0, \nu_\alpha)$ and stability parameter $\alpha_i, (i = 1, ..., n)$.
The mean first exit time of the solution process $\boldsymbol{X}_t$ satisfies the following nonlocal integral-differential boundary problem:

$$\begin{cases} Au = -1\\ u|_{D^c} = 0\end{cases} \tag{3.3}$$

where equation (3.3) describes the mean first exit time of the process from a region D to its residual set D^c, and A is the generator of the solution process $\boldsymbol{X}_t$, which can be get from Corollary A.7 in Appendix A.

$$\begin{aligned}Af = {}&\boldsymbol{b}\cdot\nabla f + \frac{1}{2}Tr[\boldsymbol{\sigma}\boldsymbol{\sigma}^T\boldsymbol{H}(f)] + \sum_{i=1}^{n}(\sigma_{L,i})^{\alpha_i}\\ &\int_{R^n\setminus\{0\}}[f(x_1,x_2,...,x_{i-1},x_i+\mathrm{d}x,x_{i+1},...,x_n) - f(x_1,x_2,...,x_n)]\nu_{\alpha_i}(\mathrm{d}x).\end{aligned} \tag{3.4}$$

3.2. Calculation and analysis of the spiking frequency

In this section, in order to describe the average spiking frequency of the Class 1 excitable Ermentrout–Kopell neural model equation (3.1), we utilize the finite difference method to calculate the mean first exit time over the interval $[-\pi/2, \theta^+]$, where θ^+ is the threshold potential described in Section 2.5. Additionally, for the sake of clarity, we provide the derivation and error estimation of the finite difference scheme in Appendix B for reference.

Figure 3 describes the relationship between the mean first exit time and the stability parameter α in a system driven solely by Lévy noise. We can find that the decline in stability parameter α results in a notable increase in

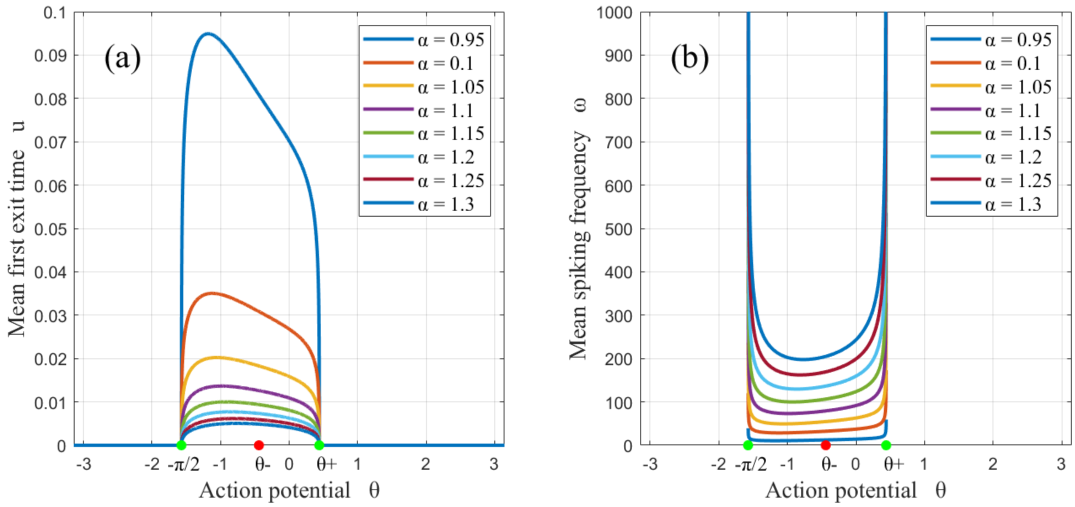


FIGURE 3. The mean first exit time (a) and spiking frequency (b) for different stability parameter α, which is calculated for the cases solely driven by Lévy noise, *i.e.* $\sigma_{\rm B} = 0, \sigma_L = 1$.

the mean first exit time, and concomitantly, a sharp decrease in the spiking frequency, thus leading the neurons into an inactive state.

Nonetheless, pure jump Lévy noise may inadequate to capture the initial potential fluctuations of neurons in a broader context. In this study, we explore a system driven by both Gaussian white noise and Lévy noise. We specifically examine the correlation between the mean first exit time and the noise intensity σ_B & σ_L across various stability parameters α. The outcomes are illustrated in Figure 4.

Figure 4 depicts the rest state θ^- as the observed value across diverse conditions, and presents the trend in the mean first exit time and spiking frequency concerning noise intensities σ_B & σ_L. It is observed that the mean first exit time generally declines as noise intensity increases. However, counterintuitively, with a small stability parameter α, a minor Gaussian white noise intensity is adequate to notably reduce the mean first exit time. We call this phenomenon the noise coupling effect. When this effect manifests, the neuronal system can substantially decrease the mean first exit time and sharply increase the spiking frequency at the cost of lower noise intensities σ_B & σ_L, thus making the neurons be in an active state. The mechanism of the noise coupling effect will be elaborated detaily in the next chapter of discussion.

4. Discussion

In this study, a data-driven approach is introduced to extract neuron models from noised data. Two models of Class 1 excitability, *i.e.* Ermentrout–Kopell neural system and FHN neuron system, are examined under the potential fluctuation of both Gaussian white noise and Lévy noise. The findings demonstrate that the noise statistics can be accurately determined, leading to a relatively precise identification of the neuron models, which allows us to go one step further and better analyze the underlying dynamic behavior of neuronal systems.

Furthermore, two key indicators, namely the mean first exit time and spiking frequency, are utilized to investigate the transitions of neurons from a resting state to an active state. In order to calculate the mean first exit time, we develop a finite difference method for solving non-local boundary value problems under Lévy noise excitation. The results reveal a noise coupling effect, *i.e.*, in a system where Gaussian white noise and Lévy noise act simultaneously, it is observed that with a small stability parameter α, a significant reduction in the mean first exit time can be achieved with an minor intensity of Gaussian white noise, rather than adopting the conventional approach of simultaneously increasing the intensities of both Gaussian white noise and Lévy noise.

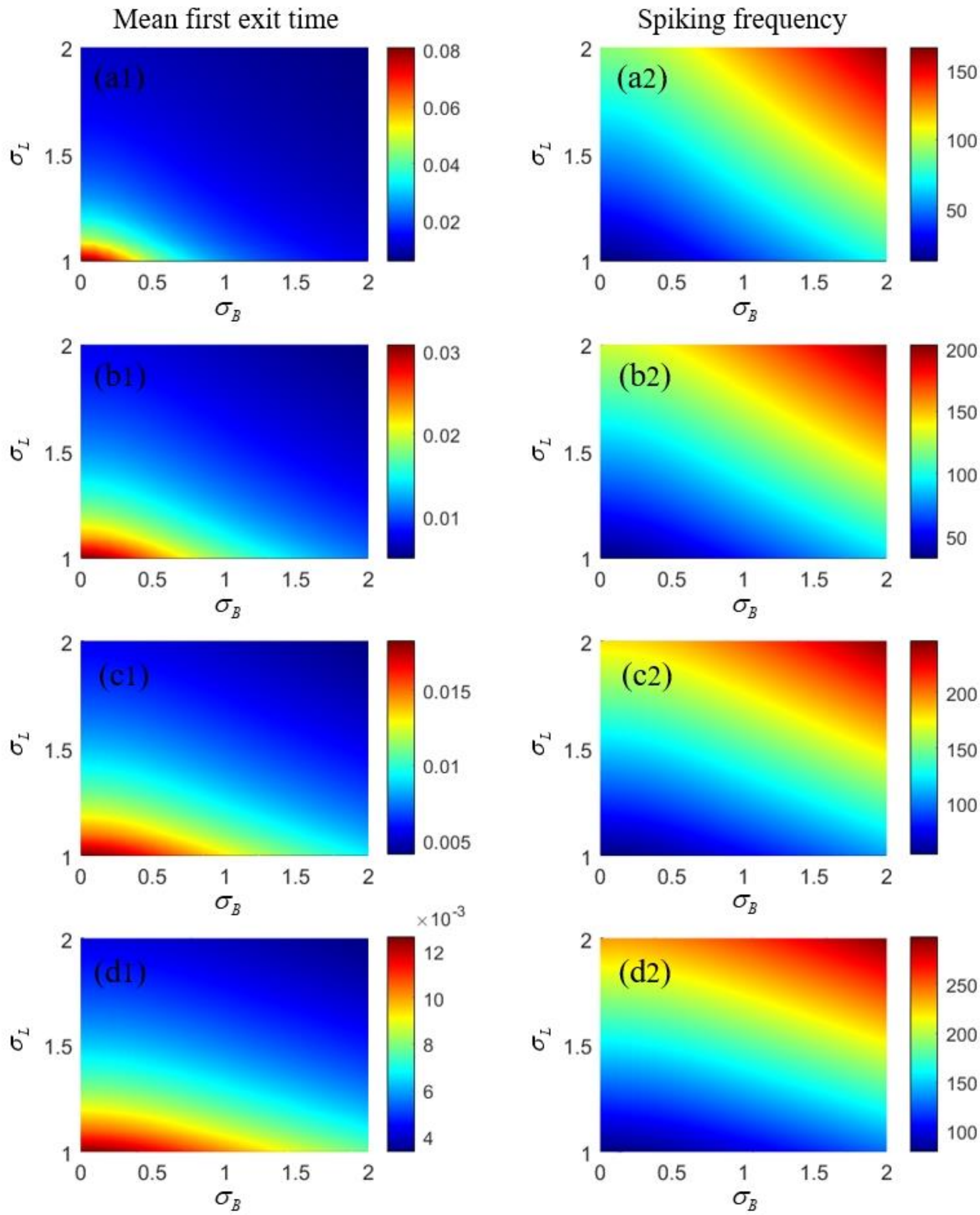


FIGURE 4. Mean first exit time and spiking frequency for different stability parameter α. Where (a): $\alpha = 0.95$, (b): $\alpha = 1$, (c): $\alpha = 1.1$, (d): $\alpha = 1.5$.

The fundamental mechanism of the noise coupling effect illustrates that, instead of simply amplifying noise intensities by multiplying the amplitude of noise-induced fluctuations, the heavy-tail property of the Lévy process (*i.e.*, the decay rate of probability density is polynomial) and the light-tail property of the Gaussian process (*i.e.*, the decay rate of probability density is exponential, which is faster than the polynomial decay rate) complement each other, promoting the exit of a dynamical system. Specifically, under the condition of moderate noise intensity, the width of the amplitude probability density distribution of the noise term holds greater significance than doubling the noise fluctuation amplitude. The noise coupling effect serves as a compelling explanation for conflicting outcomes observed by different researchers in prior studies [53–56] when analyzing the dominant factor of the exit issue in stochastic dynamics. Consequently, further exploration of the noise coupling effect holds substantial importance in addressing exit problems within stochastic dynamics.

Appendix A. Lévy process

Definition A.1. A Lévy process, $\boldsymbol{L}_t\, or\, \boldsymbol{L}(t)$, is a stochastic process satisfying the following conditions:

(i) $\boldsymbol{L}_0 = 0, a.s.$
(ii) Independent increments: For $t_1{<}t_2{<}\cdots{<}t_{n-1}{<}t_n$, the random variables $\boldsymbol{L}_{t_2} - \boldsymbol{L}_{t_1}, \cdots, \boldsymbol{L}_{t_n} - \boldsymbol{L}_{t_{n-1}}$ are independent.
(iii) Stationary increments: $\boldsymbol{L}_t - \boldsymbol{L}_s$ and $\boldsymbol{L}_{t-s}$ have the same distribution.
(iv) Stochastically continuous sample paths: For all $\delta > 0$ and all $s{\geq}0$, $P(|\boldsymbol{L}_t - \boldsymbol{L}_s| > \delta) \to 0$ as $t \to s$.

Definition A.2. (Lévy–Khintchine formula [57]) If $\boldsymbol{L}_t$ is a Lévy process with triplet $(\boldsymbol{b}, \boldsymbol{Q}, \nu)$ in R^n, then its characteristic function is:

$$\Phi_t(\boldsymbol{u}) = E\mathrm{e}^{i<\boldsymbol{u},\boldsymbol{L}_t>} = \mathrm{e}^{t\eta(\boldsymbol{u})}\, foreach t{\geq}0, \mathrm{u} \in R^n,$$

where $\eta(\boldsymbol{u}) = i\boldsymbol{b} \cdot \boldsymbol{u} - \frac{1}{2}\boldsymbol{u} \cdot \boldsymbol{Q}\boldsymbol{u} + \int_{R^n\backslash\{\mathbf{0}\}}[\mathrm{e}^{i\boldsymbol{u}\cdot\boldsymbol{y}} - 1 - iI_{\{||\mathrm{y}||<1\}}\mathrm{u}\cdot\mathrm{y}]\nu(\mathrm{d}\mathbf{y})$,
$\eta(\boldsymbol{u})$is defined for a vector $\boldsymbol{b} \in R^n$, a non-negative definite symmetric matrix $\boldsymbol{Q} \in R^{n\times n}$, and a Borel measure ν on $R^n\backslash\{\mathbf{0}\}$ for which $\int_{R^n\backslash\{\mathbf{0}\}}(||y||^2{\wedge}1)\nu(\mathrm{d}\boldsymbol{y}) < \infty$.

Definition A.3. α-stable Lévy process is a special but important class of Lévy processes, which is defined by stable random variable with the distribution $S_\alpha(\delta, \beta, \lambda)$ [57, 58]. Where $\alpha \in (0, 2]$ represents the stability parameter, $\delta \in (0, \infty)$ is the scaling parameter, $\beta \in [-1, 1]$ is the skewness parameter, and $\lambda \in (-\infty, \infty)$ denotes the shift parameter.
Specifically, the characteristic function of a rotationally symmetric α-stable Lévy process is:

$$\Phi_t(u) = E\mathrm{e}^{<u,L_t>} = \exp(-t|u|^\alpha)$$

where the jump measure is given by:

$$\nu_\alpha(\mathrm{d}\boldsymbol{y}) = c(n, \alpha)|\mathrm{y}|^{-(n+\alpha)}\mathrm{dy}$$

with the constant:

$$c(n, \alpha) = \frac{\alpha\Gamma((n + \alpha)/2)}{2^{1-\alpha}\pi^{n/2}\Gamma(1 - \alpha/2)}.$$

For a small value of $\alpha(0 < \alpha < 1)$, α-stable Lévy process exhibits a higher probability of experiencing large jumps. Conversely, for a large value of $\alpha(1 < \alpha < 2)$, the α-stable Lévy process tends to have a higher probability of small jumps. The special case is $\alpha = 2$, which corresponds to the Brown motion. More details about Lévy process can be found in references [57, 58].

Theorem A.4. *The generator of a Lévy process with triplet* $(\boldsymbol{b}, \boldsymbol{Q}, \nu)$ *is:*

$$A\varphi = \boldsymbol{b} \cdot \nabla\varphi + \frac{1}{2}Tr(\boldsymbol{Q}H(\varphi)) + \int_{R^n\backslash\{\mathbf{0}\}}[\varphi(\boldsymbol{x} + \boldsymbol{y}) - \varphi(\boldsymbol{x}) - I_{\{||\mathrm{y}||<1\}}\mathrm{y} \cdot \nabla\varphi(\mathrm{x})]\nu(\mathrm{dy})$$

where H is the Hessian matrix, and I denotes the indicative function.

To facilitate the calculation of the mean first exit time and escape probability distribution, we derive the generator of εL_t, where ε is the intensity of Lévy process.

Corollary A.5. *For Lévy process $L_t \in R^n$ with triplet $(\boldsymbol{b}, \boldsymbol{Q}, \nu)$, the generator of εL_t is:*

$$A\varphi = \varepsilon \boldsymbol{b} \cdot \nabla\varphi + \frac{\varepsilon^2}{2} Tr(\boldsymbol{Q} H(\varphi)) + \int_{R^n\setminus\{\mathbf{0}\}} [\varphi(\boldsymbol{x} + \varepsilon \boldsymbol{y}) - \varphi(\boldsymbol{x}) - \varepsilon I_{\{||\mathrm{y}||<1\}} \mathrm{y} \cdot \nabla\varphi(\mathrm{x})] \nu(\mathrm{dy}).$$

Proof:
From Definition A.2, for $L_t \in R^n$ with triplet $(\boldsymbol{b}, \boldsymbol{Q}, \nu)$, the characteristic function is given as:

$$\Phi_t(\boldsymbol{u}) = E\mathrm{e}^{i<\boldsymbol{u}, \boldsymbol{L}_t>} = \exp\{i\boldsymbol{b} \cdot \boldsymbol{u}t - \frac{1}{2} u \cdot Qut + t \int_{R^n\setminus\{\mathbf{0}\}} [\mathrm{e}^{iu\cdot y} - 1 - iI_{\{||\mathrm{y}||<1\}} \mathrm{u} \cdot \mathrm{y}] \nu(\mathrm{d}\mathbf{y})\}$$

The characteristic function of εL_t is:

$$\begin{aligned}
&\Phi'(\boldsymbol{u}) = E\mathrm{e}^{i<\boldsymbol{u}, \varepsilon \boldsymbol{L}_t>} = E\mathrm{e}^{i<\varepsilon \boldsymbol{u}, \boldsymbol{L}_t>} \\
&= \exp\{\varepsilon i\boldsymbol{b} \cdot \boldsymbol{u}t - \tfrac{\varepsilon^2}{2} \boldsymbol{u} \cdot \boldsymbol{Q}\boldsymbol{u}t + t \textstyle\int_{R^n\setminus\{\mathbf{0}\}} [\mathrm{e}^{i\varepsilon \boldsymbol{u}\cdot \boldsymbol{y}} - 1 - \varepsilon i I_{\{||\mathrm{y}||\le 1\}} \mathrm{u} \cdot \mathrm{y}] \nu(\mathrm{dy})\} \\
&= \exp\{\varepsilon i\mathbf{b} \cdot \mathbf{u}t - \varepsilon it\mathbf{u} \cdot \textstyle\int_{R^n\setminus\{\mathbf{0}\}} I_{\{||\mathbf{y}||\le 1\}} \mathbf{y} \nu(\mathrm{d}\mathbf{y}) - \tfrac{\varepsilon^2}{2} \mathbf{u} \cdot \mathbf{Q}\mathbf{u}t + t \textstyle\int_{R^n\setminus\{\mathbf{0}\}} [\mathrm{e}^{i\mathbf{u}\cdot \mathbf{y}} - 1] \nu(d(\tfrac{\mathbf{y}}{\varepsilon}))\}
\end{aligned}$$

$$\begin{aligned}
&= \exp\{\varepsilon i\boldsymbol{b} \cdot \boldsymbol{u}t - \varepsilon it\boldsymbol{u} \cdot \textstyle\int_{R^n\setminus\{\mathbf{0}\}} I_{\{||\mathrm{y}||\le 1\}} \mathrm{y} \nu(\mathrm{dy}) + it\mathbf{u} \cdot \textstyle\int_{R^n\setminus\{\mathbf{0}\}} I_{\{||\mathbf{y}||\le 1\}} \mathbf{y} \nu(d(\tfrac{\mathbf{y}}{\varepsilon})) - \tfrac{\varepsilon^2}{2} \mathbf{u} \cdot \mathbf{Q}\mathbf{u}t \\
&+ t \textstyle\int_{R^n\setminus\{\mathbf{0}\}} [\mathrm{e}^{i\mathbf{u}\cdot \mathbf{y}} - 1 - iI_{\{||\mathbf{y}||\le 1\}} \mathbf{u} \cdot \mathbf{y}] \nu(d(\tfrac{\mathbf{y}}{\varepsilon}))\} \\
&= \exp\{i[\varepsilon \mathbf{b} - \varepsilon \textstyle\int_{R^n\setminus\{\mathbf{0}\}} I_{\{||\mathbf{y}||\le 1\}} \mathbf{y} \nu(\mathrm{d}\mathbf{y}) + \textstyle\int_{R^n\setminus\{\mathbf{0}\}} I_{\{||\mathbf{y}||\le 1\}} \mathbf{y} \nu(d(\tfrac{\mathbf{y}}{\varepsilon}))] \cdot \mathbf{u}t - \tfrac{1}{2} \mathbf{u} \cdot (\varepsilon^2 \mathbf{Q}) \mathbf{u}t \\
&+ t \textstyle\int_{R^n\setminus\{\mathbf{0}\}} [\mathrm{e}^{i\mathbf{u}\cdot \mathbf{y}} - 1 - iI_{\{||\mathbf{y}||\le 1\}} \mathbf{u} \cdot \mathbf{y}] \nu(d(\tfrac{\mathbf{y}}{\varepsilon}))\}
\end{aligned}$$

Therefore, εL_t is a Lévy process with triplet:

$$(\varepsilon \boldsymbol{b} - \varepsilon \int_{R^n\setminus\{\mathbf{0}\}} I_{\{||\mathrm{y}||\le 1\}} \mathrm{y} \nu(\mathrm{dy}) + \int_{R^n\setminus\{\mathbf{0}\}} I_{\{||\mathbf{y}||\le 1\}} \mathbf{y} \nu(d(\frac{\mathbf{y}}{\varepsilon})), \varepsilon^2 \mathbf{Q}, \nu(d(\frac{\mathbf{y}}{\varepsilon})))$$

From Theorem A.4, we know the generator of εL_t is given as Corollary A.5.

Corollary A.6. *For a symmetric α-stable Lévy process ${L_t}^\alpha \in R^n$ with triplet $(\boldsymbol{b}, \boldsymbol{Q}, \nu_\alpha)$, the generator of $\varepsilon {L_t}^\alpha$ is given as:*

$$A^\alpha \varphi = \varepsilon \boldsymbol{b} \cdot \nabla\varphi + \frac{\varepsilon^2}{2} Tr(\boldsymbol{Q} H(\varphi)) + \varepsilon^{n-1+\alpha} \int_{R^n\setminus\{\mathbf{0}\}} [\varphi(\boldsymbol{x} + \boldsymbol{y}) - \varphi(\boldsymbol{x}) - I_{\{||\mathrm{y}||<1\}} \mathrm{y} \cdot \nabla\varphi(\mathrm{x})] \nu_\alpha(\mathrm{dy})$$

where $\nu_\alpha(\mathrm{dy}) = c(n, \alpha) \cdot \frac{\mathrm{d}y}{||y||^{n+\alpha}}$, $c(n, \alpha) = \frac{\alpha \Gamma((n+\alpha)/2)}{2^{1-\alpha} \pi^{n/2} \Gamma(1-\alpha/2)}$.

Proof. For α-stable Lévy process,

$$\nu_\alpha(d(\frac{\boldsymbol{y}}{\varepsilon})) = c(n, \alpha) \cdot \frac{d(\frac{\boldsymbol{y}}{\varepsilon})}{||\frac{\mathrm{y}}{\varepsilon}||^{n+\alpha}} = \varepsilon^{n-1+\alpha} \cdot c(n, \alpha) \cdot \frac{\mathrm{dy}}{||\mathrm{y}||^{n+\alpha}} = \varepsilon^{n-1+\alpha} \cdot \nu_\alpha(\mathrm{dy})$$

From Corollary A.5, the triplet corresponding to $\varepsilon {L_t}^\alpha$ is:

$$(\varepsilon \boldsymbol{b} - \varepsilon \int_{R^n\setminus\{\mathbf{0}\}} I_{\{||\mathrm{y}||\le 1\}} \mathrm{y} \nu(\mathrm{dy}) + \int_{R^n\setminus\{\mathbf{0}\}} I_{\{||\mathbf{y}||\le 1\}} \mathbf{y} \nu(d(\frac{\mathbf{y}}{\varepsilon})), \varepsilon^2 \mathbf{Q}, \varepsilon^{n-1+\alpha} \cdot \nu_\alpha(\mathrm{d}\mathbf{y}))$$

From Theorem 1, we know the generator of $\varepsilon {L_t}^\alpha$ is given as Corollary A.6. ∎

Corollary A.7. *For SDE in R^n of the following form:*

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{b}(\boldsymbol{X}_t) dt + \boldsymbol{\sigma}_B(\boldsymbol{X}_t) \mathrm{d}\boldsymbol{B}_t + \boldsymbol{\sigma}_L \mathrm{d}{\boldsymbol{L}_t}^{\boldsymbol{\alpha}}$$

where $\boldsymbol{X_t} \in R^n, \boldsymbol{b}(\boldsymbol{X_t}) \in R^n, \boldsymbol{B_t} \in R^m, \boldsymbol{\sigma}_B(\boldsymbol{X_t}) \in R^{n\times m}, \boldsymbol{\sigma_L} = diag(\sigma_{L,1}, \sigma_{L,2}, ..., \sigma_{L,n})$*is the Lévy noise intensity,* $\boldsymbol{L_t}^{\boldsymbol{\alpha}} = ({L_t}^{\alpha_1}, {L_t}^{\alpha_2}, ..., {L_t}^{\alpha_n})^T$*, and each component* ${L_t}^{\alpha_i}, (i = 1, ..., n)$ *is independent Lévy process with stability parameter* α_i.
Generator of the solution process $\boldsymbol{X_t}$ *is given as follows:*

$$Af = \boldsymbol{b} \cdot \nabla f + \frac{1}{2} Tr[\boldsymbol{\sigma}\boldsymbol{\sigma}^T \boldsymbol{H}(f)] + \sum_{i=1}^{n} (\sigma_{L,i})^{\alpha_i}$$
$$\int_{R^n \setminus \{0\}} [f(x_1, x_2, ..., x_{i-1}, x_i + \mathrm{d}x, x_{i+1}, ..., x_n) - f(x_1, x_2, ..., x_n)] \nu_{\alpha_i}(\mathrm{d}x).$$

The cumulative term in the latter half of the formula arises from the independence of each component ${L_t}^{\alpha_i}, (i = 1, ..., n)$*. For more details and more general cases of multiplicative Lévy noise, readers can refer to References. [57,59].*

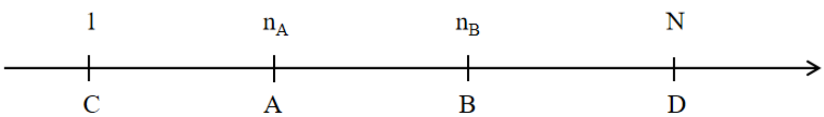
Appendix B. Difference scheme derivation and error estimation

In this section, we drive a second-order difference scheme for the stochastic dynamical system of the general form:

$$dX_t = b(X_t)dt + \sigma_B(X_t)\mathrm{d}B_t + \sigma_L dL_t^{\alpha}$$

where $X_t \in R^1$, L_t^{α} is a pure jump α-stable Lévy process with intensity σ_L. From Corollary A.7 in Appendix A, we know the generator of the solution process X_t has the following form:

$$Af(x) = b(x) \cdot \frac{\partial f}{\partial x} + \frac{1}{2}\sigma_B^2(x) \cdot \frac{\partial^2 f}{\partial x^2} + \sigma_L^{\alpha} \cdot c(1,\alpha) \cdot \int_{R^1/\{0\}} \frac{f(x+y) - f(x)}{||y||^{1+\alpha}} \mathrm{d}y$$



Let us consider the following exit problem. Given two intervals that satisfy $[\mathrm{C}, \mathrm{D}] \supset [\mathrm{A}, \mathrm{B}]$, calculate the mean first exit time from $[\mathrm{A}, \mathrm{B}]$ to the residual set $[\mathrm{C}, \mathrm{A}] \bigcup [\mathrm{B}, \mathrm{D}]$.

Firstly, we divide the interval $[\mathrm{C}, \mathrm{D}]$ into $\mathrm{N} - 1$ equal parts with length h, which lead to N nodes as described in the above diagram.

Secondly, the central difference scheme is employed to approximate the partial derivatives in generator A, which leads to the accuracy of $\mathrm{O}(\mathrm{h}^2)$. For the nonlocal boundary problem of the mean first exit time:

$$\begin{cases} Au = -1 \\ u|_{D^c} = 0 \end{cases}$$

where we employ:

$$\begin{cases} \dfrac{\partial u}{\partial x} = \dfrac{u_{i+1} - u_{i-1}}{2h} \\ \dfrac{\partial^2 u}{\partial x^2} = \dfrac{u_{i+1} - 2u_i + u_{i-1}}{h^2} \end{cases} \quad for\, i \in [n_A + 1, n_B - 1]$$

for the non-local integral term in the generator, we have:

TABLE C.1. Identified Lévy motion for the FHN neuron system.

The stability parameter α			Lévy noise intensity σ_L		
True	Learn	Error	True	Learn	Error
2	2.0311	1.6×10^{-2}	2	2.0721	3.6×10^{-2}
1.5	1.4927	4.9×10^{-3}	2	2.0337	1.7×10^{-2}
1	1.0093	9.3×10^{-3}	2	1.9896	5.2×10^{-3}
0.5	0.5015	3.1×10^{-3}	2	2.0591	2.9×10^{-2}
0.25	0.2538	1.5×10^{-2}	2	1.9913	4.3×10^{-3}

$$\sigma_L^\alpha \cdot c(1,\alpha) \cdot \int_{R^1/\{0\}} \frac{u(x+y)-u(x)}{||y||^{1+\alpha}} \mathrm{d}y = \sigma_L^\alpha \cdot c(1,\alpha) \cdot [\int_{-\infty}^{A-x} \frac{u(x+y)-u(x)}{||y||^{1+\alpha}} \mathrm{d}y + \int_{A-x}^{B-x} \frac{u(x+y)-u(x)}{||y||^{1+\alpha}} \mathrm{d}y + \int_{B-x}^{+\infty} \frac{u(x+y)-u(x)}{||y||^{1+\alpha}} \mathrm{d}y]$$

since the mean first exit time $u(x)$ vanishes outside the region $[A, B]$, the above equation can be simplified as:

$$\begin{aligned} &\sigma_L^\alpha \cdot c(1,\alpha) \cdot [\int_{-\infty}^{A-x} \frac{-u(x)}{||y||^{1+\alpha}} \mathrm{d}y + \int_{A-x}^{B-x} \frac{u(x+y)-u(x)}{||y||^{1+\alpha}} \mathrm{d}y + \int_{B-x}^{+\infty} \frac{-u(x)}{||y||^{1+\alpha}} \mathrm{d}y] \\ &= \sigma_L^\alpha \cdot c(1,\alpha) \cdot \{-\frac{u(x)}{\alpha}[\frac{1}{(x-A)^\alpha} + \frac{1}{(B-x)^\alpha}] + \int_{A-x}^{B-x} \frac{u(x+y)-u(x)}{||y||^{1+\alpha}} \mathrm{d}y\} \end{aligned}$$

which can be discretized as:

$$\begin{aligned} &\sigma_L^\alpha \cdot c(1,\alpha) \cdot \{-\frac{u(x)}{\alpha}[\frac{1}{(x-A)^\alpha} + \frac{1}{(B-x)^\alpha}] + \int_{A-x}^{B-x} \frac{u(x+y)-u(x)}{||y||^{1+\alpha}} \mathrm{d}y\} \\ &= \sigma_L^\alpha \cdot c(1,\alpha) \cdot \{-\frac{u(x_i)}{\alpha}[\frac{1}{(x_i-A)^\alpha} + \frac{1}{(B-x_i)^\alpha}] + \sum_{j=n_A-i}^{n_B-i} \frac{u(x_{i+j})-u(x_i)}{||x_{i+j}-x_i||^{1+\alpha}}\} \\ &for i \in [n_A+1, n_B-1] \end{aligned}$$

Leading order error of the above quadrature rule is $-\zeta(\alpha-1)u^{''}(x)h^{2-\alpha} + o(h^2)$[60], where ζ is the Riemann-zeta function.

Therefore, the equation that governs the mean first exit time can be discretized as the following form:

$$\begin{cases} b(x_i) \cdot \frac{u_{i+1}-u_{i-1}}{2h} + \frac{1}{2}\sigma_B^2(x_i) \cdot \frac{u_{i+1}-2u_i+u_{i-1}}{h^2} - \sigma_L^\alpha \cdot c(1,\alpha) \cdot \frac{u_i}{\alpha} \cdot [\frac{1}{(x_i-A)^\alpha} + \frac{1}{(B-x_i)^\alpha}] & \\ + \sigma_L^\alpha \cdot c(1,\alpha) \cdot \sum_{j=n_A-i}^{n_B-i} \frac{u_{i+j}-u_i}{||x_{i+j}-x_i||^{1+\alpha}} = -1 & for\ i \in [n_A+1, n_B-1] \\ u_i = 0 & for\ i \in [1, n_A] \bigcup [n_B, n] \end{cases}$$

Finally, we solve the above linear system utilizing the LU factorization or the Krylov subspace iterative method GMRES.

APPENDIX C. SYSTEM IDENTIFICATION FOR THE FHN NEURON SYSTEM

To clarify the structure of this paper, the identification results of the FHN neuron system equation (2.16) are listed separately in this section, where the sampling interval $h = 0.01$, and the function dictionary consists of polynomial bases up to the third order.
where the procedures used for system identification in this section are openly available in Github [53].

TABLE C.2. Identified drift term $b_1(\boldsymbol{x})$ for different stability parameter α, where $\sigma_B = 1, \sigma_L = 2$.

Basis	True	$\alpha = 0.5$		$\alpha = 1$		$\alpha = 2$	
		Learn	Error	Learn	Error	Learn	Error
1	0	0		0		0	
x	1	1.0891	8.9×10^{-2}	1.0472	4.7×10^{-2}	0.9981	1.9×10^{-3}
y	–1	-0.9893	1.1×10^{-2}	–1.0829	8.3×10^{-2}	–1.0257	2.6×10^{-2}
x^2	0	0		0		0	
xy	0	0		0		0	
y^2	0	0		0		0	
x^3	–1/3	–0.3417	2.5×10^{-2}	–0.3401	2.1×10^{-2}	–0.3395	1.7×10^{-2}
x^2y	0	0		0		0	
xy^2	0	0		0		0	
y^3	0	0		0		0	

TABLE C.3. Identified drift term $b_2(\boldsymbol{x})$ for different stability parameter α, where $\sigma_B = 1, \sigma_L = 2$.

Basis	True	$\alpha = 0.5$		$\alpha = 1$		$\alpha = 2$	
		Learn	Error	Learn	Error	Learn	Error
1	0.24	0.2513	4.7×10^{-2}	0.2486	3.6×10^{-2}	0.2451	2.1×10^{-2}
x	0.2	0.2103	5.1×10^{-2}	0.2085	4.2×10^{-2}	0.2047	2.6×10^{-2}
y	0	0		0		0	
x^2	0	0		0		0	
xy	0	0		0		0	
y^2	0	0		0		0	

TABLE C.4. Identified diffusion term $a_{11}(\boldsymbol{x})$ for different stability parameter α, where $\sigma_B = 1, \sigma_L = 2$.

Basis	True	$\alpha = 0.5$		$\alpha = 1$		$\alpha = 2$	
		Learn	Error	Learn	Error	Learn	Error
1	2	2.0849	4.2×10^{-2}	2.0722	3.6×10^{-2}	1.9847	7.7×10^{-3}
x	0	0		0		0	
y	2	1.9892	5.4×10^{-3}	2.0638	3.2×10^{-2}	2.0276	1.4×10^{-2}
x^2	0	0		0		0	
xy	0	0		0		0	
y^2	1	1.0528	5.3×10^{-2}	1.0641	6.4×10^{-2}	1.0182	1.9×10^{-2}

TABLE C.5. Identified diffusion term $a_{12}(\boldsymbol{x}) = a_{21}(\boldsymbol{x})$ for different stability parameter α.

Basis	True	$\alpha = 0.5$		$\alpha = 1$		$\alpha = 2$	
		Learn	Error	Learn	Error	Learn	Error
1	0	0		0		0	
x	1	1.0627	6.3×10^{-2}	1.0322	3.2×10^{-2}	0.9925	7.5×10^{-3}
y	0	0		0		0	
x^2	0	0		0		0	
xy	0	0		0		0	
y^2	0	0		0		0	

TABLE C.6. Identified diffusion term $a_{22}(\boldsymbol{x})$ for different stability parameter α.

Basis	True	$\alpha = 0.5$		$\alpha = 1$		$\alpha = 2$	
		Learn	Error	Learn	Error	Learn	Error
1	0	0		0		0	
x	0	0		0		0	
y	0	0		0		0	
x^2	1	0.9782	2.2×10^{-2}	1.0337	3.4×10^{-2}	0.9918	8.2×10^{-3}
xy	0	0		0		0	
y^2	0	0		0		0	

Funding

This research is supported by the National Natural Science Foundation of China (No. 12172167).

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability statement

The data and procedures that support the findings of this study are openly available in Github [61].

References

[1] A.L. Hodgkin, The local electric changes associated with repetitive action in a non-medullated axon. *J. Physiol.* **107** (1948) 165–181.

[2] J. Rinzel and G.B. Ermentrout, Analysis of neural excitability and oscillations. *Methods Neuronal Model.* (1989) 251–291.

[3] E.M. Izhikevich, Neural excitability, spiking and bursting. *Int. J. Bifurc. Chaos* **10** (2000) 1171–1266.

[4] F.C. Hoppensteadt and E.M. Izhikevich, Weakly Connected Neural Networks. Springer (1997).

[5] G.B. Ermentrout, Type I membranes, phase resetting curves, and synchrony. *Neural Comput.* **8** (1996) 979–1001.

[6] S.A. Prescott, Excitability: types I, II, and III. *Encyclop. Computat. Neurosci.* **1** (2014) 1–7.

[7] R.G. Xie, W.G. Chu, S.J. Hu and L. Ceng, Characterization of different types of excitability in large somatosensory neurons and its plastic changes in pathological pain states. *Int. J. Mol. Sci.* **19** (2018) 161–174.

[8] E.M. Izhikevich, Dynamical Systems in Neuroscience. Massachusetts Institute of Technology (2007).

[9] L. Holden and T. Erneux, Slow passage through a Hopf bifurcation: form oscillatory to steady state solutions. *SIAMJ. Appl. Math.* **53** (1993) 1045–1058.

[10] L. Holden and T. Erneux, Understanding bursting oscillations as periodic slow passages through bifurcation and limit points. *J. Math. Biol.* **31** (1993) 351–365.

[11] Y.A. Bedrov, G.N. Akoev and O.E. Dick, Partition of the Hodgkin–Huxley type model parameter space into regions of qualitatively different solutions. *Biol. Cybernet.* **66** (1992) 413–418.

[12] A. Majumdar, J. Ockendon, P. Howell and E. Surovyatkina, Transitions through critical temperatures in nematic liquid crystals. *Phys. Rev. E* **88** (2013) 022501.

[13] D. Premraj, K. Suresh, T. Banerjee and K. Thamilmaran, An experimental study of slow passage through Hopf and pitchfork bifurcations in a parametrically driven nonlinear oscillator. *Commun. Nonlinear Sci. Numer. Simul.* (2016) S1007570416000105.

[14] J. Rinzel and G. Huguet, Nonlinear dynamics of neuronal excitability, oscillations, and coincidence detection. *Commun. Pure Appl. Math.* **66** (2013) 1464–1494.

[15] G. Huguet, X. Meng and J. Rinzel, Phasic firing and coincidence detection by subthreshold negative feedback: divisive or subtractive or, better, both. *Front. Computat. Neurosci.* **11** (2017) 00003.

[16] D.W. Van, Modeling neural activity. *ISRN Biomath.* (2013) 1–37.

[17] A. Shilnikov, Complete dynamical analysis of a neuron model. *Depart. Math. Statist. Georgia State Univ.* **68** (2012) 305–328.

[18] T.B. Luke, E. Barreto and P. So, Complete classification of the macroscopic behavior of a heterogeneous network of theta neurons. *Neural Computat.* **25** (2013) 3207–3234.

[19] T.C. Redman, Data Driven: Profiting from Your Most Important Business Asset. Harvard Business School Publishing (2008).

[20] S.L. Brunton and J.N. Kutz, Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control. Cambridge University Press (2019).

[21] W. Aalst, Data Science in Action. Springer Nature (2016).

[22] S.L. Brunton, J. Proctor and J. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Natl. Acad. Sci.* **113** (2016) 3932–3937.

[23] K. Champion, B. Lusch, J. Kutz and S.L. Brunton, Data-driven discovery of coordinates and governing equations. *Natl. Acad. Sci.* **116** (2019) 22445–22451.

[24] B. Tzen and M. Raginsky, Neural stochastic differential equations: deep latent Gaussian models in the diffusion limit. (2019) arxiv preprint arXiv:1905.09883.

[25] R.T. Chen, Y. Rubanova, J. Bettencourt and D.K. Duvenaud, Neural ordinary differential equations. *Conf. Neural Inform. Process. Syst.* (2018).

[26] C.A. García, P. Félix, J. Presedo and D.G. Márquez, Nonparametric estimation of stochastic differential equations with sparse Gaussian processes. *Phys. Rev. E* **96** (2017) 022104.

[27] P. Batz, A. Ruttor and M. Opper, Approximate Bayes learning of stochastic differential equations. *Phys. Rev. E* **98** (2018) 022109.

[28] Y. Li and J. Duan, Extracting governing laws from sample path data of non-Gaussian stochastic dynamical systems. *J. Statist. Phys.* **180** (2022) 30.

[29] Y. Li and J. Duan, A data-driven approach for discovering stochastic dynamical systems with non-Gaussian Lévy noise. *Physica D* **417** (2021) 132830.

[30] R. Reza and M. Tabar, Analysis and Data-based Reconstruction of Complex Nonlinear Dynamical Systems. Springer, New York, NY (2019).

[31] S. Klus, F. Nüske, S. Peitz, J.H. Niemann, C. Clementi and C. Schütte, Data-driven approximation of the Koopman generator: model reduction, system identification, and control. *Physica D* **406** (2020) 132416.

[32] Y. Li and J. Duan, Discovering transition phenomena from data of stochastic dynamical systems with Lèvy noise. *Chaos* **30** (2020) 093110.

[33] N. Takeishi, Y. Kawahara and T. Yairi, Subspace dynamic mode decomposition for stochastic Koopman analysis. *Phys. Rev. E* **96** (2017) 033310.

[34] X. Chen, L. Yang, D. Jian-Qiao and G.E. Karniadakis, Solving inverse stochastic problems from discrete particle observations using the Fokker–Planck equation and physics-informed neural networks. *SIAM J. Sci. Comput.* **43** (2021).

[35] L. Yang, C. Daskalakis and G.E. Karniadakis, Generative ensemble regression: learning particle dynamics from observations of ensembles with physics-informed deep generative models. *SIAM J. Sci. Comput.* **44** (2022).
[36] G.B. Ermentrout and N. Kopell, Multiple pulse interactions and averaging in systems of coupled neural oscillators. *J. Math. Biol.* **29** (1991) 195–217.
[37] R.E. DeVille and V.E. Eric, A Nontrivial Scaling Limit for Multiscale Markov Chains. *J. Statist. Phys.* **126** (2007) 75–94.
[38] R.E. Deville and E. Vanden, Self-induced stochastic resonance for Brownian ratchets under load. *Commun. Math. Sci.* **5** (2007) 431–466.
[39] M.E. Yamakou and J. Jost, Coherent neural oscillations induced by weak synaptic noise. *Nonlinear Dyn.* **93** (2018) 2121–2144.
[40] M.E. Yamakou and J. Jost, Control of coherence resonance by self-induced stochastic resonance in a multiplex neural network. *Phys. Rev. E* **100** (2019) 022313.
[41] C. Mead, Neuromorphic electronic systems. *Proc. IEEE* **78** (1990) 1629–1636.
[42] C.L. Nikias and M. Shao, Signal Processing with Alpha-Stable Distributions and Applications. Wiley, New York (1995).
[43] A. Dave, Stable non-Gaussian random processes by G. Samorodnitsky and M.S. Taqqu. *Math. Gazette* **79** (1995) 625.
[44] F. Zhan and S. Liu, Response of electrical activity in an improved neuron model under electromagnetic radiation and noise. *Front. Computat. Neurosci.* **11** (2017) 107–118.
[45] E. Marius, Yamakou and D.T. Tat, Lévy noise-induced self-induced stochastic resonance in a memristive neuron. *Nonlinear Dyn.* **107** (2022) 2847–2865.
[46] G.B. Ermentrout, Type I membranes, phase resetting curves, and synchrony. *Neural Comput.* **8** (1996) 979–1001.
[47] J. Guckenheimer, R. Harris-Warrick, J. Peck and A. Willms, Bifurcations, bursting and spike frequency adaptation. *J. Comput. Neurosci.* **4** (1997) 257–277.
[48] H.C. Tuckwell, Random perturbations of the reduced Fitzhugh–Nagumo equation. *Physica Scripta* **46** (1992) 481–484.
[49] B. Lindner and G.L. Schimansky, Analytical approach to the stochastic FitzHugh–Nagumo system and coherence resonance. *Phys. Rev. E* **60** (1999) 7270–7276.
[50] I. Franović, K. Todorović, M. Perc, N. Vasović and N. Burić, Activation process in excitable systems with multiple noise sources: One and two interacting units. *Phys. Rev. E* **92** (2015).
[51] J.A. Reinoso, M.C. Torrent and C. Masoller, Emergence of spike correlations in periodically forced excitable systems. *Phys. Rev. E* **94** (2016).
[52] A.C. Scott, The electrophysics of a nerve fiber. *Rev. Mod. Phys.* **47** (1975) 487–533.
[53] R. Cai, X.L. Chen, J.Q. Duan *et al.*, Lévy noise-induced escape in an excitable system. *J. Statist. Mech. Theory Exp.* **1** (2017) 063503.
[54] A. Tesfay, D. Tesfay, A. Khalaf and J. Brannan, Mean exit time and escape probability for the stochastic logistic growth model with multiplicative Levy noise. *Stochast. Dyn.* (2021) 2150016.
[55] T. Gao and J. Duan, Quantifying model uncertainty in dynamical systems driven by non-Gaussian Lévy stable noise with observations on mean exit time or escape probability. *Commun. Nonlinear Sci. Numer. Simul.* **39** (2016) 1–6.
[56] J. Ren, C. Li, T. Gao, X. Kan and J. Duan, Mean exit time and escape probability for a tumor growth system under non-Gaussian noise. *Int. J. Bifurc. Chaos* **22** (2012) 1250090.
[57] D. Applebaum, Lévy Process and Stochastic Calculus, 2nd edn. Cambridge University Press, Cambridge (2009).
[58] D. Jian-Qiao, An Introduction to Stochastic Dynamics. Science Press, Beijing (2015).
[59] X. Sun, X. Li and Y. Zheng, Governing equations for probability densities of Marcus stochastic differential equations with Lévy noise. *Stochast. Dyn.* (2016) 1750033.
[60] A. Sidi and M. Israeli, Quadrature methods for periodic singular and weakly singular Fredholm integral equations. *J. Sci. Comput.* **3** (1988) 201–231.
[61] https://github.com/Douglas-catty/Neuronal-System